

## Twitter

Készítette a Közigazgatási és Igazságügyi minisztérium E-közigazgatási  
Szabad Szoftver Kompetencia Központja  
Budapest, 2013



A projekt az Európai Unió támogatásával, az Európai Regionális Fejlesztési Alap társfinanszírozásával valósul meg.

Kódszám: EKOP–1.2.15

Ez a Mű a Creative Commons Nevezd meg! – Így add tovább! 3.0 Unported  
Licenc feltételeinek megfelelően szabadon felhasználható.

A dokumentum legfrissebb változata letölthető a honlapunkról:

<http://szabadszoftver.kormany.hu/>

## Tartalomjegyzék

A Twitter informatikai infrastruktúrája.....	2
Egy tweet útja a felhasználótól az olvasóig.....	2
A Twitter és a nyílt forrású fejlesztés.....	3

## Twitter

A Twitter az egyik legnagyobb online közösségi oldal és mikroblogging szolgáltatás több mint 500 millió aktív felhasználóval. A felhasználók az oldalon maximum 140 karakteres szöveges üzeneteket tehetnek közzé, amelyeket a felhasználót követők elolvashatnak. A Naponta 300-400 millió tweet üzenetet küldenek, és 1,6 milliárd keresési kérés fut be az oldalt kiszolgáló szerverekre,<sup>1</sup> amely csak hatékony és megbízható informatikai infrastruktúrával kezelhető. A Twitter kezdettől fogva nyílt forrású megoldásokra építkezett, a cég sikere egyben a nyílt forráskód sikerének a bizonyítéka.

## A Twitter informatikai infrastruktúrája

Chris Aniszczyk, a Twitter nyílt forrással foglalkozó vezetője, a LinuxCon észak-amerikai technológiai konferencián beszélt arról 2012 augusztusában, hogy milyen nyílt forrású megoldásokat használnak a cégnél. Kiemelte, hogy a „Twitter filozófiája, hogy minden nyílt forrású legyen.”<sup>2</sup> Az alap operációs rendszer a Linux 2.6.39 kernelre épül, míg az adatbázisok kezelésére a MySQL adatbázist használják. A cég egy saját MySQL verziót fejleszt, a szintén nyílt megoldásaikat pedig megosztják a MySQL fejlesztőivel. Az elemzések és keresések elvégzésére a Cassandra adatbázist, a Hadoop adatfeldolgozó keretrendszert, a Lucene keresőtechnológiát, a Pig magas szintű adatelemző platformot és számos egyéb Apache fejlesztést használnak. A Twitter támogatja az Apache Alapítványt, és saját fejlesztéseiket itt is igyekeznek visszajuttatni az eredeti fejlesztőkhöz. Az egyre növekvő forgalom kezelésében a Memcached alkalmazás áll rendelkezésükre, amelyet nagy mértékben használnak a gyorsítótár infrastruktúrában. A Memcached segítségével felgyorsíthatják a dinamikus webalkalmazásokat az adatbázis-terhelés csökkentésével.<sup>3</sup>

## Egy tweet útja a felhasználótól az olvasóig<sup>4</sup>

A Twitter évente körülbelül 2.8 milliárd üzenetet kezel, amely átlagosan 5000 üzenetet jelent másodpercenként, de természetesen nem egyenletes terhelés mellett. Először minden tweet regisztrálásra kerül, majd pedig egy egyedi azonosítót kap a snowflake program használatával. Ezután a geolocation (helymeghatározási) adatokat dolgozza fel a Rockdove. Ezt követően az üzeneteket ellenőrzik egy URL-rövidítő és spam-figyelő, a t.co segítségével. Miután ezek megtörténtek, az üzenet bekerül a MySQL adatbázisba a Gizzard alkalmazáson keresztül, amely egy rugalmas keretrendszer konzisztens elosztott adatbázisok létrehozására. Ezután jelzi csak a felhasználó webböngészője felé a Twitter, hogy az üzenet rendben elküldésre került.

<sup>1</sup> <http://en.wikipedia.org/wiki/Twitter>

<sup>2</sup> <http://www.zdnet.com/how-twitter-tweets-your-tweets-with-open-source-7000003526/>

<sup>3</sup> <http://techcrunch.com/2012/08/30/how-twitter-uses-open-source/>

<sup>4</sup> <http://www.zdnet.com/how-twitter-tweets-your-tweets-with-open-source-7000003526/>

Ekkor azonban még nem került ki a tweet a felhasználó üzenetfolyamába. Először a keresők számára teszik elérhető a Firehose alkalmazásfejlesztői felületen (API) keresztül. Ezután áll készen az üzenet arra, hogy kikerüljön a világ elé. Ezt a folyamatot a FlockDB kezeli, amely egy nyílt forrású gráfadatbázis, amely a Gizzardon keresztül szerzi meg az adatokat a MySQL adatbázisból. A FlockDB tárolja a Twitter felhasználóit és a kapcsolatokat a felhasználók között. Miután tehát kiderült, hogy kik a felhasználó követői a Twitteren, végül kikerül az üzenet a megfelelő beállításokkal a felhasználó oldalára.

Az itt leírt folyamat 350 milliszekundumot vesz igénybe, és jól láthatjuk, hogy minden Twitter üzenet elküldésével számos nyílt forrású alkalmazást vesz igénybe a felhasználó.

## A Twitter és a nyílt forrású fejlesztés

A Twitter nemcsak felhasználja és továbbfejleszti az elérhető nyílt forrású alkalmazásokat, hanem saját maga is fejleszt több nyílt forrású megoldást is. A Iago egy terhelésgenerátor, amellyel az üzembe helyezés előtt tudják tesztelni a szolgáltatásaikat. A Zipkin egy elosztott követőrendszer, amelyet azért alkotott meg a cég, hogy az időzítési adatokat össze tudják gyűjteni a szolgáltatások számára, ezzel a Twitter gyorsabbá válik. A Scalding pedig egy Scala könyvtár, amely lehetővé teszi számukra, hogy egyszerűen írjanak MapReduce folyamatokat a Hadoop keretrendszerben.

A Twitter által fejlesztett alkalmazások legnagyobb hányadát az Apache License Version 2.0 licenc alatt adják ki. A dokumentációkat pedig a Creative Commons Attribution 3.0 Unported licenc alatt teszik közzé. Ennek eredményeképpen bárki szabadon felhasználhatja, módosíthatja és továbbadhatja a dokumentációkat, forráskódokat és a nyílt forrású projektjeikben futó példaprogramokat mindaddig, amíg betartják a licencfeltételeket.

Annak érdekében, hogy a Twitter hivatalosan is a nyílt forrás támogatója legyen, 2012 augusztusában csatlakozott a cég a Linux Alapítványhoz ezüst fokozatú támogatóként.<sup>5</sup> A Linux Alapítvány végzi a Linux kernel fejlesztését, a munkájukat számos olyan vállalat támogatja, amelyek az informatikai struktúrájukat nyílt forrásra építették. A Linux Alapítvány ügyvezető igazgatója szerint, ez a lépés „mutatja, milyen fontossá válik a Linux az ilyen hatalmas méretű internetes vállalatok számára, amelyek lényegében Linuxon futnak, és szinte a teljes infrastruktúrájukat nyílt forrású szoftvekre építik fel”.

<sup>5</sup> <http://arstechnica.com/business/2012/08/twitter-joins-linux-foundations-fight-for-open-source-software/>